CFP 1239 EP

(19) Federal Republic of Germany - German Patent Office

(12) Patent Specification

(10) DE 195 28 425 C1

(51) Int. Cl.: H 04 N 13/00

(21) Application Number: 195 28 425.9-31

(22) Date of Filing: 2nd August 1995

(43) Date of Publication: --

(45) Date of Publication of Issue of Patent: 15th May 1996

Appeals may be lodged within three months of publication

---

(73) Assignee:

Siemens AG, 80333 Munich, Germany

(72) Inventors:

Riegel, Thomas, Dipl.-Inform, 81739 Munich, Germany;

Ziegler, Manfred, Dipl.-Inform, 81739 Munich, Germany

(56) For the assessment of patentability the following documents were taken into consideration:

DE 40 27 471 C2

EP 06 50 301 A2

---

(54) Configuration and procedure for the automatic selection of two camera signals from multiple camera signals by determining the position of the head of a viewer.

(57). A head position (Pj) of a viewer (B) is determined by a head-following unit (KVE), and two camera signals (Ki; Ki + x) are selected, and requested from a transmitter unit (SE). The requested cameras (Kai; Kai + x) are switched through to a stereoscopic encoding unit (SC) by a switching unit (S), where they are encoded into an encoded video data stream (CV). This encoded video data stream (CV) is transmitted to a receiver unit (EE), where it is decoded and presented to the viewer (B) on a stereoscopic screen (DIS).

In this way, a very simple, automated selection of two necessary camera signals (Ki; Ki + x) out of multiple camera signals (Ki) is achieved.

Description

During the transmission of so-called multiview stereo picture sequences, in which more than two cameras record a scene, several problems exist during the transmission of the camera signals that describe the recorded scene. With multiview stereo picture sequences, the receiver needs two camera signals that describe the same scene from different views. A depth effect is achieved for the viewer through the existence of a disparity between the pictures of the recorded camera signals, that is to say, through the local displacement of corresponding pixels between the pictures of the two camera signals, which are presented via a receiver on a stereoscopic screen. In other words, the viewer has the impression that he or she is observing a three-dimensional picture.

If a scene is recorded by several cameras, all camera signals were transmitted by all cameras to the receiver, and the selection of the two necessary camera signals were to take place in the receiver, then an extremely large transmission capacity would be required, which would include much unnecessary redundancy.

As only two camera signals are needed in the receiver unit (the two 'correct' ones for the viewer), it is desirable to actually transmit only the necessary two camera signals. Furthermore, it is desirable to be able to

vary the view of the scene without driving up the cost.

The use of more than two cameras for stereoscopic picture recording is hitherto known from EP 0650301 A2. Furthermore, DE 40 27 471 C2 concerns a spatial representation of pictures that depends on the detected position of the viewer.

Therefore, the object of the invention is to solve the problem of specifying a configuration and a procedure by which two camera signals are chosen in a simple manner from camera signals produced by a plurality of n cameras which record a scene from different views, the signals being chosen automatically by a viewer of the scene, who is located at the receiver.

This problem is solved by the configuration according to Claim 1, and the procedure according to Claim 3.

One advantage of the configuration according to the present invention lies, above all, in the fact that the configuration is very simple and therefore only minor enhancements to configurations are necessary, which are in any case required in order to realize a stereoscopic video data stream on the basis of two connected cameras which record a scene.

Also, the procedure according to the present invention is characterized by its simplicity, which leads to a very fast execution of the procedure, whereby it is possible to

implement the procedure in real time, without delay effects, which may be disturbing for the viewer, when selecting another view of an observed scene.

The use of the procedure according to Claim 4 for exception handling in case of a mistake, and therefore as a so-called fallback solution, in the use of a model-based encoding procedure for the encoding of a video data stream, significantly increases the dependability of the model-based encoding procedure.

Further refinements of the invention have resulted from the attached claims.

A preferred embodiment of the invention and an advantageous application of the procedure according to the invention are presented in the drawings and are more closely described in the following.

The following are shown:

Fig. 1 is a diagram that shows the configuration of a receiver unit, a transmitter unit, and multiple cameras, which are selected based on the position of the viewer's head;

Fig. 2 is a flow chart that describes the procedure according to the invention in accordance with Claim 3; and

Fig. 3 is a flow chart that shows the use of the procedure according to the invention as a fallback solution in model-based encoding.

Based on Figs. 1 to 3, the configuration according to the invention, the procedure according to the invention, and an application of the procedure according to the invention are further illustrated.

In Fig. 1, the configuration according to the invention is described, which features at least the following components:

- a transmitter unit SE which features at least the following components:
  - a plurality of cameras Kai, which record a scene from several views, in other words from different perspectives; a first index, i, which unambiguously identifies each camera, is a number that lies in the range between 1 and n, the values 1 and n being contained in the range;
  - a switching unit S, with n inputs, wherein each input of switching unit S is connected to a separate camera Kai, and two outputs,
  - a stereoscopic encoding unit SC for encoding a video data stream CV to be transmitted;
- a receiver unit EE which features at least the following components:
  - a stereoscopic decoder unit SD, in which the received encoded video data stream CV is decoded,
  - a head-following unit KVE, in which the head

position Pj of a viewer B is determined, and in which the necessary two views are selected from the determined coordinates of the head position, in other words, two necessary camera signals Ki are selected, wherein, as well as the first index i denoting the camera signal provided by the i-th camera, a second index j, which unambiguously identifies each head position Pj of the viewer B, is a number lying in the range between 1 and n - 1, the values 1 and n - 1 being contained in the range;

- a stereoscopic screen DIS, which presents a decoded video data stream DCV to the viewer B in such a way as to create a stereoscopic effect for the viewer B, that is to say, an impression of depth is formed as the picture is watched.

The cameras Kai record a scene SZ. The cameras Kai are arranged in any desired manner.

This is a prerequisite for the fact that in any desired selection of two cameras Kai and Kai + x, in which the number x lies in the range between 1 and n - 1, a stereoscopic effect can be achieved for the viewer B.

The transmitter unit SE features at least the n cameras Kai, the switching unit S, and the stereoscopic encoding unit SC.

The switching unit S is connected by its two outputs to the stereoscopic encoding unit SC. The design of the stereoscopic encoding unit SC is well-known to one skilled in the art and is, for example, described in (R. ter Horst, A Demonstrator for Coding and Transmission of Stereoscopic Video Signals, Proceedings of 4th European Workshop on Three-dimensional Television, Rome, 20-21 October 1993, P. 273-280, 1993).

The video data stream which is to be transmitted, and which is encoded in the stereoscopic encoding unit SC, is transmitted to the receiver unit EE. Each of the inputs of the switching unit S is connected to a separate camera Kai.

Corresponding to a camera request message KAN, which will be further described, the cameras Kai and Kai + 1, which are requested by the camera request message KAN, are selected in the switching unit S. In this way, each of the requested camera signals Ki and Ki + 1, from the cameras Kai and Kai + 1, are encoded and transmitted.

Thus in each case only the requested camera signals are selected and switched through 3 to the stereoscopic encoding unit SC, where they are encoded 4 into the video data stream CV and transmitted to the receiver unit EE.

In the receiver unit EE, the encoded video data stream CV is received and decoded 5. Subsequently, the decoded video data stream is presented on the stereoscopic screen

DIS to the viewer B.

In 6, the receiver unit EE features at least the stereoscopic decoder unit SD, the head-following unit KVE, and the stereoscopic screen DIS.

The stereoscopic decoder unit SD is connected to the stereoscopic screen DIS. In the head-following unit KVE, the position of the head of the viewer B is determined by means of a further camera, with which the head of the viewer B is recorded and detected, and the spatial location of the head of the viewer B, that is the head position Pj of the viewer B, is determined.

Based on the head position Pj, wherein the second index j lies in the range between 1 and n - 1, the camera signals Ki and Ki + 1, which are assigned to the head position Pj, are selected and requested by the head-following unit KVE.

In the head-following unit KVE, the head position Pj of the viewer B is determined 1 and the horizontal movements of the viewer B with respect to the screen are followed. For this, at least one camera that records the head of the viewer B is used. The head-following unit KVE may operate, for example, according to the procedure described in the patent application with application number 19516664.7.

In the procedure described there, the spatial location

is determined for a recorded target object, which is classified during the procedure described there. This procedure is used in the procedure according to the present invention so that the head of the viewer B may be detected and the movements of the head followed. In this way the procedure described in patent application 19516664.7 determines the spatial location of the head of the viewer B of the stereoscopic screen DIS.

However, only those coordinates that describe horizontal movements are used in the procedure according to the present invention. Neither of the other spatial coordinates are of interest for the procedure according to the invention and for this reason are not required.

In Fig. 2, the procedure in which the two necessary camera signals Ki and Ki + 1 are selected is described.

First of all, both of the acceptable extreme positions for the head position Pj of the viewer B are defined, that is to say the head position P1 and the head position Pn - 1.

In the case that a realistic transmission is desired, it is advantageous to set up the cameras Kai equidistant from each eye, which in this case means that the distance between the head position P1 and the head position Pn - 1 would be the same as the distance between the camera Ka1 and the camera Kan - 1. The necessary camera signals can then be selected and transmitted, depending on the position

of the viewer B.

If the head of the viewer is located at head position P1, in other words on the far left of the configuration described in Fig. 1, or even further to the left, both of the camera signals K1 and K2 are selected, that is to say the camera signals from the two cameras positioned at the left Ka1 and Ka2. These signals are then requested by the head-following unit KVE.

This set-up also provides for a situation in which camera signals K1 and K2 are not selected, but rather there could for example be a camera positioned on the far left Ka1 and a further camera, Ka1 + x, and the camera signals produced by these cameras K1 and K1 + x are selected. The selection of the cameras is always dependent on the configuration of the stereoscopic screen DIS.

If the head of the viewer B is located at head position Pn - 1 or to the right of it, the head-following unit KVE assigns the camera signals Kn - 1 and Kn to the head position Pn - 1, and requests these camera signals from the transmitter unit SE.

The set-up also provides for a situation in which camera signals Kn - 1 and Kn are not selected, but rather there could for example be a camera positioned on the far left Ka1 and a further camera, Ka 1 + x, and the camera signals produced by these cameras Kn and Kn - x are

selected. In this case too, the selection of the cameras is always dependent on the configuration of the stereoscopic screen DIS.

If the head of the viewer B is located between the discrete head positions Pj, the head position Pj of the viewer B that is assigned is the one that is closest to the actual head position of the viewer B.

In general, therefore, the camera signal Ki as a left picture and the camera signal Ki + x as a right picture for a head position Pj of the viewer are offered through the head-following unit KVE.

It is important to define the n - 1 head positions Pj of the viewer B independently from the camera arrangement. This is carried out depending on the structure of the stereoscopic screen DIS and from the intended effects for the viewer B.

This means that two new camera signals are requested by the head-following unit KVE. In order to improve and/or exaggerate the stereoscopic effect, that is the impression of depth for the viewer B, it is also possible to increase the space between the cameras.

If the stereoscopic effect is supposed to be reduced, naturally the space between the cameras can also be reduced.

After the head position is determined 1 by the head-following unit KVE, as previously described, the head-

following unit KVE selects 2 the necessary two views, that is to say, the necessary camera signals Ki and Ki + x, and requests these from the transmitter unit SE.

This happens in form of a camera request message KAN that is transmitted by the receiver unit EE to the transmitter unit SE.

After the transmitter unit SE has received the camera request message FAN, the switching unit S switches in such a way that both inputs of the switching unit S are connected to the requested cameras Kai and Kai + x, so that the requested camera signals Ki and Ki + x are switched through 3 to the stereoscopic encoding unit SC.

In the stereoscopic encoding unit SC, the switched through camera signals Ki and Ki + x are encoded 4 into an encoded video data stream CV to be sent.

This encoded video data stream CV is then transmitted to the receiver unit EE.

The received encoded video data stream CV is decoded in the stereoscopic decoder unit SD 5 within the receiver unit EE, and is presented 6 to the viewer B on the stereoscopic screen DIS.

In Fig. 3, an advantageous application of the procedure according to the invention is shown.

One problem in the deployment of a model-based encoding procedure MC lies in the way in which an error

situation 8 in the encoding procedure MC is handled when it occurs. In practice this means when an error occurs in the model-based encoding procedure, that is to say, when a picture was incorrectly modeled by the model-based encoding MC.

In the scope of this application, under the model-based encoding procedure MC, a procedure in which the scene SZ, having been recorded by the n cameras, is presented through three-dimensional models is to be understood. This enables a further reduction in the picture data of the recorded scene SZ.

In this case, exception handling that enables correct transmission of the video data stream for as long as is required for a correct model for the pictures of the video data stream to be produced, that is to say, calculated is necessary. The previously described procedure can be used for this exception handling.

This means that, if it is known that a model was created for a bad quality video data stream CV, the model-based encoding MC of the video data stream CV, which enables a larger compression of the picture data, is interrupted, and the procedure according to the invention is carried out 10. So the scene is recorded as appropriate by the two cameras Kai and Kai + 1, which are selected by the head position Pj of the viewer B.

If a new model that once more correctly describes the picture data were created 9, the model-based encoding MC can be carried out again 7.

In this way, the procedure according to the invention can be used in conjunction with the model-based encoding MC of a video data stream as a fallback solution of the model-based encoding MC.

Claims

1. A configuration for the automatic selection of two cameras, which produce two camera signals that together form a stereoscopic picture sequence, out of a plurality of camera signals (Ki; i = 1…n) produced by n cameras (Kai; i = 1…n):

- wherein, in a transmitter unit (SE), the n cameras (Kai), which record several views of a scene, as well as one switching unit (S) are available, and are arranged in any desired manner;
- wherein, in the transmitter unit (SE), a stereoscopic encoding unit (SC) is provided for encoding an encoded video data stream (CV) to be sent;
- wherein, in the transmitter unit (SE), the switching unit (S) is provided, wherein two outputs of the switching unit (S) are connected to the stereoscopic

encoding unit (SC), and wherein n inputs of the switching unit (S) are connected to the n cameras (Kai. . . Kan), which provide two camera signals (Ki; Ki + x) that are selected by a head-following unit (KVE) of a receiver unit (EE), the one number (x) being any number in the range between 1 and n;

- wherein, in the receiver unit (EE), the head-following unit (KVE) is provided to determine the head position (Pj; j = 1… n - 1), and to select the two camera signals (Ki; Ki + x) which are assigned to the head position (Pj);
- wherein, in the receiver unit (EE), a stereoscopic decoder unit (SD) is provided for decoding the received encoded video data stream (CV), and
- wherein, in the receiver unit (EE), a stereoscopic screen (DIS), which is connected to the output of the stereoscopic decoder unit (SD), is provided.

2. The configuration according to claim 1, wherein the distance between the n cameras (Kai) is uniform.

3. A procedure for the automatic selection of two cameras (Kai; Kai + x), which produce two camera signals (Ki; Ki + x), which together form a stereoscopic picture sequence, out of a plurality of camera signals (Ki; i = 1…n) produced by n cameras (Kai; I = 1…n):

- wherein, in a receiver unit (EE), a head position

(Pj; j = 1… n - 1) of a viewer (B) is determined by a head-following unit (KVE);

- wherein, two camera signals (Ki, Ki + 1) which are produced by two cameras (Kai; Kai + x) that are assigned to the head position (Pj) by the head-following unit (KVE) are selected, wherein a number (x) is any desired number in the range between 1 and n;

- wherein, the receiver unit (EE) sends a camera request message (KAN) to a transmitter unit (SE), by which the two necessary camera signals (Ki; Ki + x) are requested;

- wherein, the transmitter unit (SE) receives and stores the camera request message (KAN);

- wherein, in the transmitter unit (SE), the two camera signals (Ki; Ki + 1) which are selected by the head-following unit (KVE) are switched through to a stereoscopic encoding unit (SC) through a switching unit (S);

- wherein, in the transmitter unit (SE), an encoded video data stream (CV) to be sent is encoded in a stereoscopic encoding unit (SC);

- wherein, the encoded video data stream (CV) to be sent is transmitted from the transmitter unit (SE) to the receiver unit (EE);

- wherein, the received encoded video data stream (CV) is stored in the receiver unit (EE); and

- wherein, in the receiver unit (EE), the received encoded video data stream (CV) is decoded in a stereoscopic decoder unit (SD).

4. An application of the procedure according to Claim 3, wherein, in the execution of a model-based encoding procedure (MC) to encode the video data stream (CV), if the model-based encoding (MC) drops out, the procedure is executed for as long as is required until the model-based encoding (MC) can once more be implemented.

Three pages of drawings attached

Fig. 2

1. KVE determines Pj

2. Ki, Ki + x selection

3. S switches through Ki, Ki + x

4. Encoding of CV

5. Decoding of CV

6. Presentation on DIS

Time t

Fig. 3

7. MC

8. Mistake in MC occurs

9. MC is possible once again

10. Switch to invention